Data Conversion Laboratory
www.dclab.com

Data Conversion for Presentation into IADS
IADS User's Conference

May 6-8, 2003

Presented by:

Naveh Greenberg, Project Manager
ngreenberg@dclab.com
(718)357-8700 ext. 258

The SGML/XML Challenge

Your Data will be useful only if the tagging is consistent. How to effectively and efficiently produce consistently tagged data is the biggest challenge.

Parsing is not Sufficient

Just because a file parses doesn't mean that it has been correctly tagged. It just means that you haven't violated any of the DTD Rules.

Parsing is not Sufficient -- Continued

Printed Reference:

3-3. Operation of Main Switchboard.

1. With emergency power ON, set all circuit breakers to OFF.
2. Set BYPASS switch to BYPASS position.
3. Refer to paragraph 3-44.

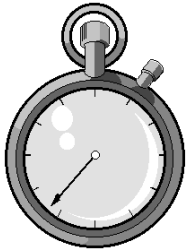
Example 1 – This will parse:

```
<para>3-3.<emphasis type="uline">Operation of Main
Switchboard.</emphasis></para>
<para>  1. With emergency power ON, set all circuit breakers
to OFF. </para>
  <para>  2. Set BYPASS switch to BYPASS position. </para>
  <para>  3. Refer to paragraph 3-44.</para>
```

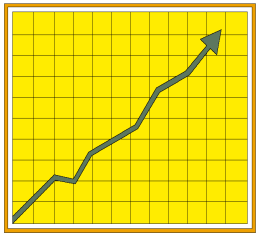
Example 2 – This will also parse and tagged correctly:

```
<proc id="para3-3">
<title>Operation of Main Switchboard.</title>
<step1><para>With emergency power ON, set all circuit breakers to
OFF.</para></step1>
  <step1><para>Set BYPASS switch to BYPASS position.</para></step1>
  <step1><para>Refer to paragraph <xref taskid="para3-
```

Conversion Objectives



- ***Data Capture Accuracy***
- ***Tagging Consistency***
- ***Time to Completion***
- ***Cost Effective***
- ***Scalability***
- ***Long Term Utility***



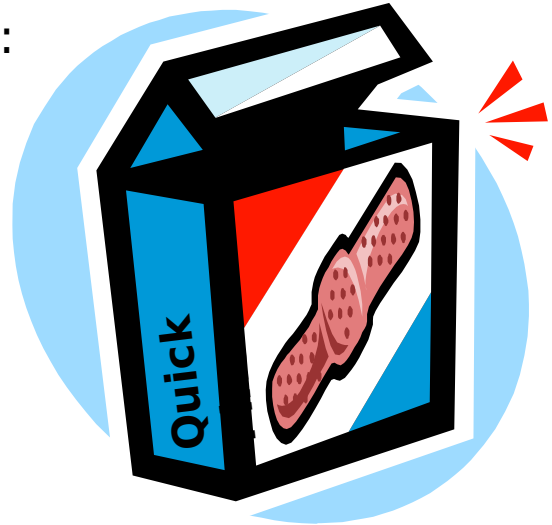
The "It's quick to fix later" trap …

If…

- 5000 page project
- 5 minutes per page to fix
- 25,000 minutes => 417 hours
- At 7 hours per day (no breaks) => 59.6 days

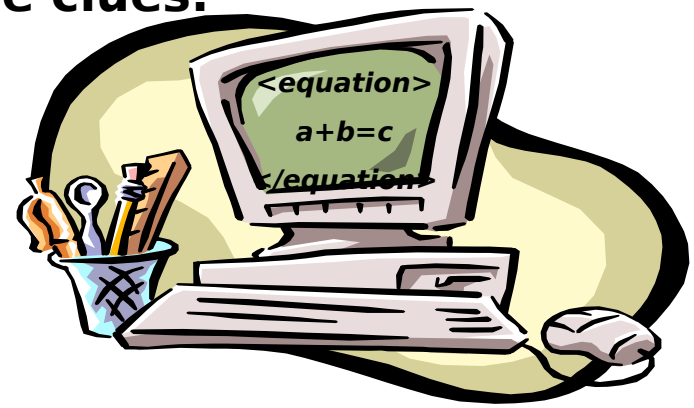
But worse…

- The tagging is inconsistent resulting in:
  - The Content/Document Management System will be useless
  - IETM won't work
  - Data won't render correctly



Keys to Consistent SGML/XML

- **Bulletproof DTD**
- **Tagging Rules have to be crystal clear**
  - **Need to be consistent**
- **Structured Tagging done by computer**
  - **Using the tagging rules**
- **Content–Level Tagging aided by computer**
  - **Pre-editing gives the software clues.**



Checklist for Analysis

- ✓ Tagging of Paragraphs, Headers, etc.
- ✓ Cross-References/Linking
- ✓ Lists
- ✓ Tables
- ✓ Graphics
- ✓ Index
- ✓ Footnotes
- ✓ Special Characters
- ✓ Math
- ✓ Table of Contents

… It's Not Just Text!

Not all Lists Are Created Equal …

Example: Numbered Lists vs. Series of Steps

- Do you need to differentiate between a simple numbered list and a series of steps?
- They appear similar to software, however, they have different meanings

Sequential List Example

Repair the following:

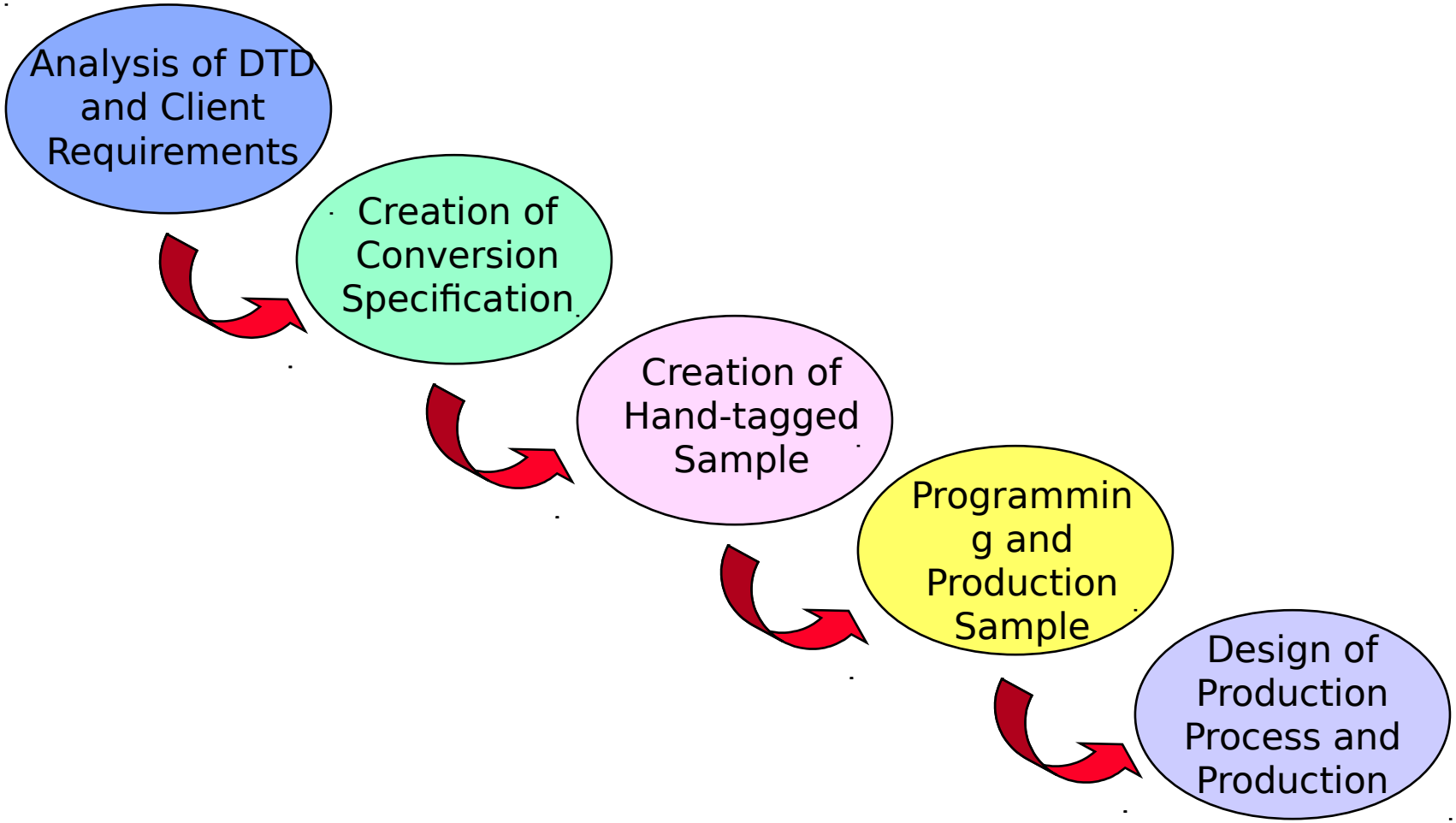
1. Flap tracks #1 and #2 in each wing.
2. Landing gear and doors.
3. Baggage door handle shaft.

Steps Example

If gust lock found engaged, proceed as follows:

1. Check control wheel and column for excessive movement.
2. Examine flight control surfaces for distortion.
3. Examine elevator stop bracket for cracks.

How is the conversion designed and developed?


Analysis of DTD and Client Requirements
Creation of Conversion Specification
Creation of Hand-tagged Sample
Programming and Production Sample
Design of Production Process and Production

Case Study - Blackhawk

- Converted approximately 20k pages of 38784 based DMWRs from Paper to 40051A/2361A based SGML. Files are viewed in IADS.

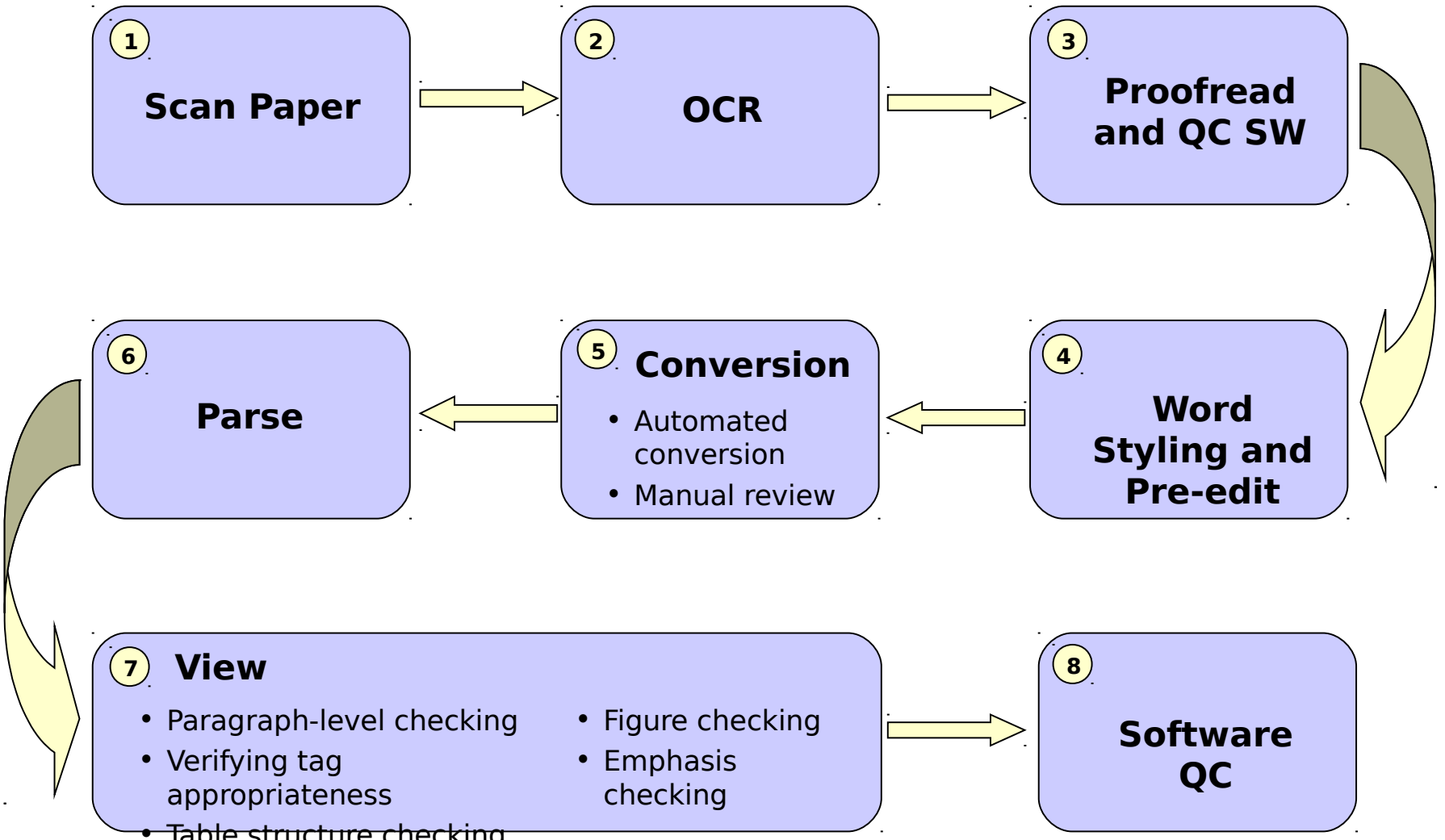
Case Study – CH47D

- Converted approximately 3k pages of TMs from Microsoft Word to 38784 based SGML. Files are viewed in IADS.

Challenges

- ***Paper***
- ***SME involvement***
- ***Inconsistent appearance***
- ***Rearranging text to fit standard***
- ***Cross references***
- ***SPEC/DTD Conflicts***
- ***Create files that are “IADS friendly”***

Challenge – Paper Source



Challenge – SME Involvement Needed

Table 2-2. AVON Preventive Maintenance Checks and Services

Step	Item to be Inspected	Task-Hours
1	EXTERIOR SURFACES. Clean the RT. Cleaning instructions are contained in Paragraph 2.10. Inspect exposed metal surfaces for rust, corrosion, and bare spots. Touchup finish or refinish equipment in accordance with Paragraph 2.15.2.	0.04
2	CABLES AND CONNECTORS. Check all cables for evidence of chafing, cracking, or excessive strain. Refer to a higher category of maintenance for replacement of defective cables. Check all electrical connectors for dents, cracks, or improper mating. Check connector cap and fuse caps for security.	0.05

SME Involvement Needed - Continued

CALS:

```
<table>
<title id="tbl2-2">AVON Preventive Maintenance Checks and
Services</title>
<tgroup cols="3">
<colspec colnum="1" colname="c1" colwidth="2.71in">
<colspec colnum="2" colname="c2" colwidth="2.44in">
<colspec colnum="3" colname="c3" colwidth="1.15in">
<thead>
<row>
<entry colname="c1" align="center">Step</entry>
<entry colname="c2" align="center">Item to be Inspected</entry>
<entry colname="c3" align="center">Task-Hours</entry>
</row>
</thead>
<tbody>
<row>
<entry colname="c1" align="center">1</entry>
<entry colname="c2" align="center">EXTERIOR SURFACES. Clean the RT.
Cleaning instructions are contained in <xref taskid="para2.10"
callout="Paragraph 2.10">.  Inspect exposed metal surfaces for rust,
corrosion, and bare spots.  Touchup finish or refinish equipment in
accordance with <xref taskid="para2.15.2" callout="Paragraph
2.15.2">.</entry>
<entry colname="c3" align="center">0.04</entry>
</row>
```

SME Involvement Needed - Continued

According to the DTD a PMCS table is actually comprised of: itemno, interval, manhours, checked, pmcsproc

CONTENT:

```
<pmcstable>
<title id="tbl2-2">Preventive Maintenance Checks and Services</title>
<pmcs-entry>
  <itemno>1</itemno>
  <interval></interval>
  <manhours>0.04</manhours>
  <checked>EXTERIOR SURFACES.</checked>
  <pmcsproc>
     <pmcsstep1>
      <para>Clean the RT.  Cleaning instructions are contained in <xref
href="para2.10"          callout="Paragraph 2.10">.  Inspect exposed metal
surfaces for rust, corrosion, and bare      spots.  Touchup finish or refinish
equipment in accordance with <xref href="para2.15.2"
callout="Paragraph 2.15.2">.</para>
    </pmcsstep1>
</pmcsproc>
</pmcs-entry>
</pmcstable>
```

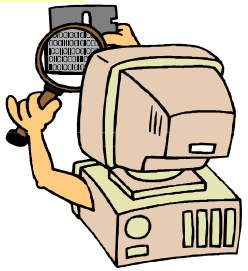
Challenge – Inconsistent Appearance

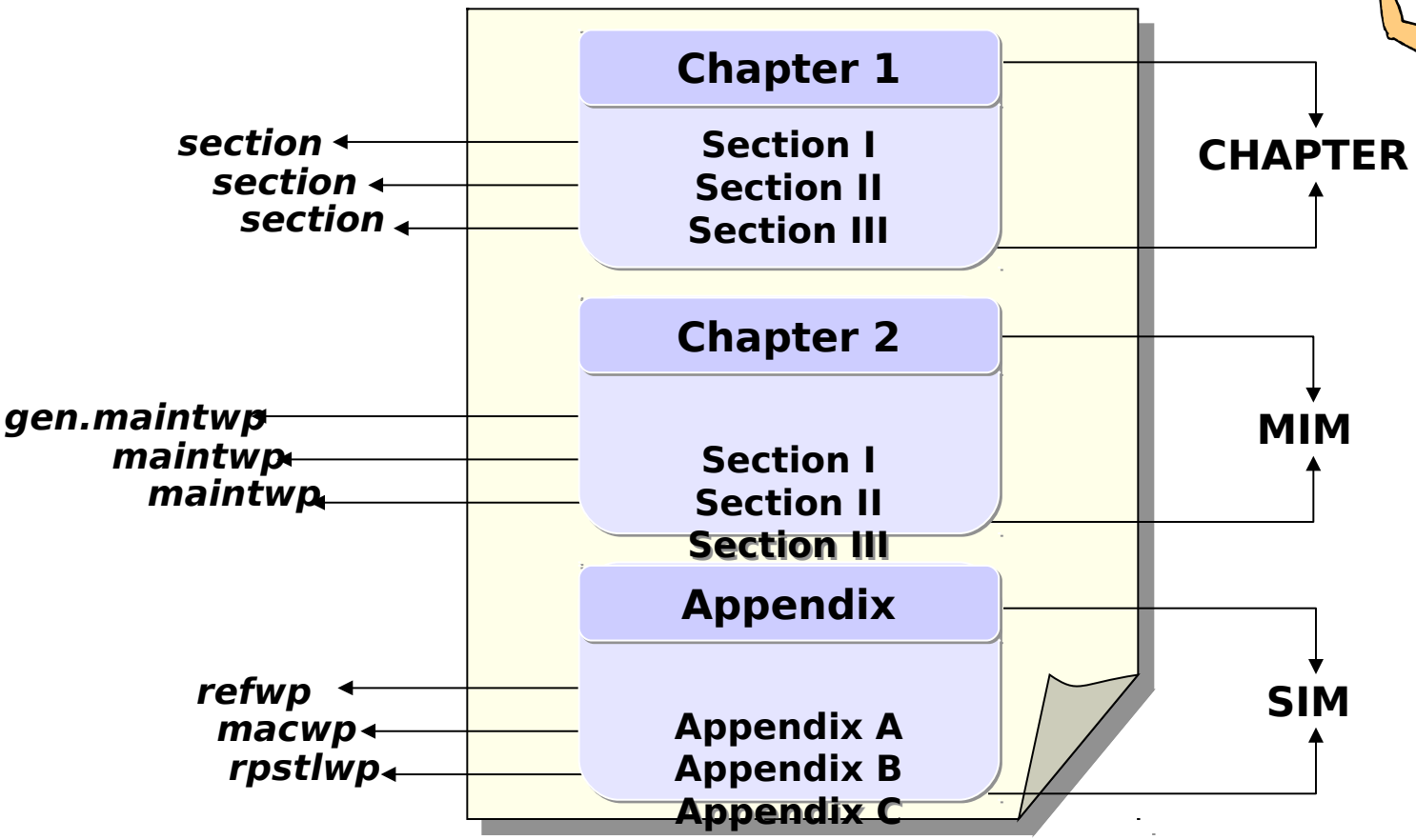
Item No.	Interval B=Before/ D=During/ A=After			Location Item to Check/ Service	Procedure	Not Fully Mission Capable if:
	B	D	A			
1				SOFT TOP SICPS		
1-1	√		√	Soft Top Assembly		
1-1.1	√		√	Soft Top Hardware Fittings	Check interior and exterior Soft Top snap fittings for damage.	

B - Before D - During A - After W - Weekly

(1) ITEM NO.	(2) INTERVAL				(3) ITEM TO BE INSPECTED PROCEDURE:	(4) EQUIPMENT IS NOT READY/AVAILABLE IF:
	B	D	A	W		
1.					S-589B/TRC-173B	
	•		•	•	a. Inspect shelter exterior for:	
					Surface area damage that could cause internal water leakage.	Holes in surface areas allow water to come in contact with equipment.

Challenge – Rearranging text to fit standard.





If it was consistent, the process would be simplified

Rearranging text to fit standard - Continued

CHAPTER 1

INTRODUCTION

Section I. General

1-1. Scope. These instructions are for use by depot/ contractor personnel. They apply to the Bench Test Set (BTS) P/N 245-476854-001, manufactured by Canadian Marconi Company, 2442 Trenton Avenue, Montreal, Quebec, Canada H3P 1Y9 and, in case of conflict, take precedence over all other documents pertinent to its overhaul and inspection.

1-2. Deviations and Exceptions.

a. Contractor Deviations. When any work segment as set forth in this depot maintenance cannot be accomplished, or can be accomplished only in a manner other than specified, prior approval of the procuring activity shall be obtained by immediately submitting to the contracting officer a written notice per DOD-STD-480 or MIL-STD-481, short form.

b. Depot Deviations. Refer to DOD-STD-480 or MIL-STD-481, short form.

1-3. Maintenance Forms and Records. Department of the Army forms and procedures used for equipment maintenance shall be those prescribed by DA Pamphlet 738-751, Functional User's Manual for the Army Maintenance Management System - Aviation (TAMMS-A).

1-4. Deficiency Reporting. Deficiency reports shall be made out and submitted to the U.S. Army Aviation Systems Command (AVSCOM), ATTN: AMSAV-QRF, 4300 Goodfellow Boulevard, St. Louis, MO 63120-1798 to suggest timely corrections and improvements to test equipment described in the DMWR. It is not necessary to show a new design or list a better way to do a procedure. Simply tell why the design is unfavorable or why a procedure is difficult. Deficiency reports will be prepared using the Standard Form 368 (SF 368), Product Quality Deficiency Report. Instructions for preparing deficiency reports are provided in DA Pamphlet 738-751. A reply will be furnished directly to you.

Rearranging text to fit standard - Continued

CHAPTER 1

INTRODUCTION

Section I. General

<SCOPE ID="PARA1-1">

~~1-1. Scope.~~ These instructions are for use by depot/ contractor personnel. They apply to the Bench Test Set (BTS) P/N 245-476854-001, manufactured by Canadian Marconi Company, 2442 Trenton Avenue, Montreal, Quebec, Canada H3P 1Y9 and, in case of conflict, take precedence over all other documents pertinent to its overhaul and inspection. </SCOPE>

<DEVIATION ID="PARA1-2">

~~1-2. Deviations and Exceptions.~~

a. Contractor Deviations. When any work segment as set forth in this depot maintenance cannot be accomplished, or can be accomplished only in a manner other than specified, prior approval of the procuring activity shall be obtained by immediately submitting to the contracting officer a written notice per DOD-STD-480 or MIL-STD-481, short form.

b. Depot Deviations. Refer to DOD-STD-480 or MIL-STD-481, short form. </DEVIATION>

<MFRR ID="PARA1-3">

~~1-3. Maintenance Forms and Records.~~ Department of the Army forms and procedures used for equipment maintenance shall be those prescribed by DA Pamphlet 738-751, Functional User's Manual for the Army Maintenance Management System - Aviation (TAMMS-A). </MFRR>

<EIR ID="PARA1-4">

~~1-4. Deficiency Reporting.~~ Deficiency reports shall be made out and submitted to the U.S. Army Aviation Systems Command (AVSCOM), ATTN: AMSAV-QRF, 4300 Goodfellow Boulevard, St. Louis, MO 63120-1798 to suggest timely corrections and improvements to test equipment described in the DMWR. It is not necessary to show a new design or list a better way to do a procedure. Simply tell why the design is unfavorable or why a procedure is difficult. Deficiency reports will be prepared using the Standard Form 368 (SF 368), Product Quality Deficiency Report. Instructions for preparing deficiency reports are provided in DA Pamphlet 738-751. A reply will be furnished directly </EIR>

Rearranging text to fit standard - Continued

According to the DTD <ginfowp> is: (wpidinfo, scope, mfrr, eir….deviation…).

When tagging Para 1-2 will be moved after Para 1-4.

```
<ginfowp wpno="G00001-55-4920-412">
<wpidinfo><maintlvl level="depot"><eicnomen><sysnomen>
<name>BENCH TEST SET VERTICAL INSTRUMENT DISPLAY SYSTEM</name>
<modelno></modelno><nsn>4920-01-112-5905</nsn><partno>245-476854-001</partno>
<eic></eic></sysnomen></eicnomen>
<title>GENERAL</title>
</wpidinfo>
<scope id="para1-1">....</scope>
<mfrr id="para1-3">&mfrr.state;....</mfrr>
<eir id="para1-4">&eir.state;....</eir>
<deviation id="para1-2">&deviation.state;.....</deviation>
```

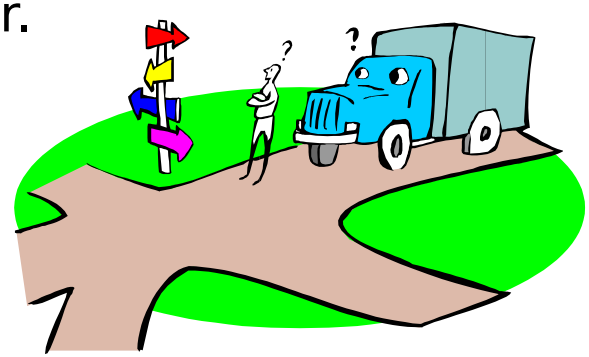
Challenge - Cross-References/Linking

Example: Automated Links

- See figure 15.5
- Refer to fig. 15.5
- See figures 15-5, 15-6 and 15-9 thru 15-13

However…

- Refer to figure below
- As illustrated on previous page
- … drawing 15.5 years at hard labor.
- See figure 15.1 in volume II of …



"Just make a left at the sign up ahead"

Challenge - Cross-References/Linking - Continued

Example: Automated Links

- See Appendix A
- Refer to paragraph 3-1
- See chapter 2.

However in 40051/2361…

- Appendix A is probably REFWP
- Paragraph 3-1 is now <geninfo> of every MAINTWP.
- Chapter 2 is now a MIM.

Challenge – SPEC/DTD Conflicts

According to the 40051A Specs

Bulleted lists are not allowed in warning, cautions, or notes.

But the 2361A DTD does allow us to use random list

<!ELEMENT (warning | caution) - - (icon-set | graphic | para | seqlist)+

-(figure | table | note) >

<!ELEMENT para – o (%content;)

<!ENTITY % content "(%text; | (%list;) | anchor | email | figure | graphic | note | proponent | table | telephone)+">

SPEC/DTD Conflicts -- Continued

WARNING

Dry Cleaning Solvent
(Stoddard Solvent)
P-D-680

- Combustible - do not use near open flames, near welding areas, or on hot surfaces.
- Prolonged contact of skin with liquid can cause dermatitis. Repeated inhalation of vapor can irritate nose and cause dizziness.
- If any liquid contacts skin or eyes, immediately flush affected area thoroughly with water. Remove solvent-saturated clothing. If vapors cause dizziness, go to fresh air.

WARNING

FLIGHT SAFETY PART

The power takeoff gearshaft is a flight safety parts. Critical characteristics are:

1) Centering/bore 1.5762/1.5765 diameter.

2) Centering/bore 1.5749/1.5754 diameter.

3) Crack limits.

Challenge – *Creating files that are "IADS friendly"*

```
<?Stylesheet href=".\styles\macsd.sty"?>
<?Title href="TM 11-5826-258-24"?>
<?embedwarnings?>
<?frame id="cover">


<?frame id="para1-1" label="1-1. "><para0 id="para1-1"
label="1-1. ">


<extref externalid="picture" xrefid="fig11-5826-258-24-fig1-1"
pretext="fig. 1-1"><?HOTSPOT id="fig. 1-1" linktype="picture"
href="./graphics/11-5826-258-24_fig1-1.CGM -T 'fig. 1-1' -W
(0,0,50,100)"><?/HOTSPOT>
```

Any
Questions?